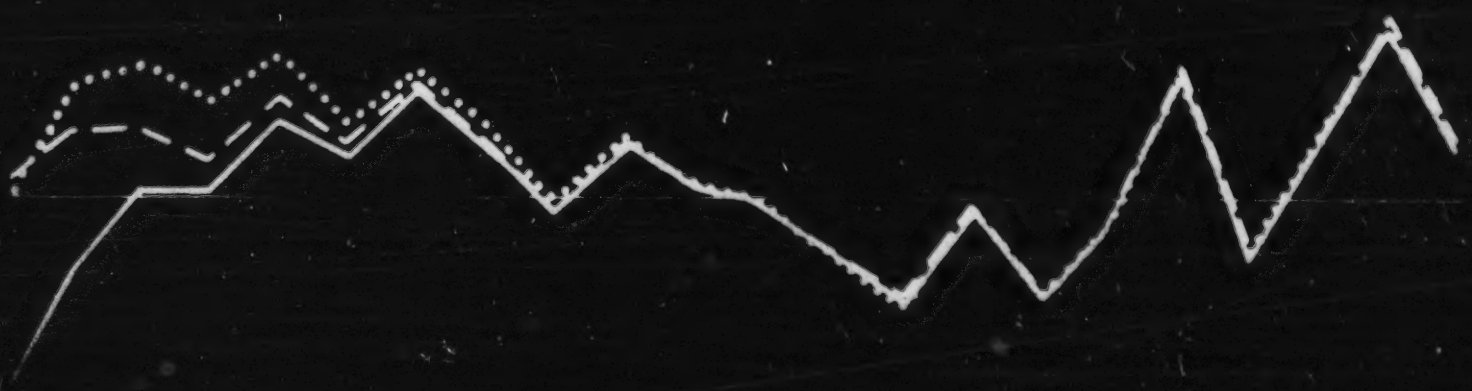
Environnement Canada
Service canadien des forêts

Environment Canada
Canadian Forestry Service

Système national d'équations pour évaluer la masse anhydre du peuplier faux-tremble *Populus tremuloides* Michx.

F. Evert

Rapport d'information PI-X-24F
Institut forestier national de Petawawa



86-05225/

INSTITUT FORESTIER NATIONAL DE PETAWAWA

L'Institut forestier national de Petawawa (IFNP) a été créé le 1er avril 1979, suite au fusionnement de la Station expérimentale forestière de Petawawa et de deux Instituts établis à Ottawa; celui de l'aménagement forestier et celui de la recherche sur les feux de forêt. En 1980, est venue s'y adjoindre la Direction de la statistique forestière et des systèmes.

Son mandat, comme celui des autres établissements du Service canadien des forêts, est de promouvoir une meilleure gestion et une utilisation plus rationnelle des ressources forestières du Canada, pour le bien économique et social de tous les Canadiens. En tant qu'Institut national, il doit s'attacher à des problèmes qui débordent le cadre régional ou qui nécessitent des compétences particulières de même qu'un équipement coûteux, non disponible aux installations régionales du Service canadien des forêts. La plupart du temps, ses recherches sont effectuées en étroite collaboration avec le personnel des centres régionaux ou des services forestiers des provinces.

L'Institut mène en parallèle deux programmes de recherche:

RECHERCHES SUR LES FEUX ET EN TÉLÉDÉTECTION. D'immenses étendues de terres forestières productives sont ravagées chaque année au Canada par les feux de forêt. Les recherches portent sur le comportement des feux de forêt, sur de nouvelles méthodes pour les combattre. Elles s'attardent aussi à l'évaluation de matériel d'extinction et de retardateurs. Elles s'intéressent encore au développement de systèmes informatisés de défense contre les incendies, systèmes que ne tardent pas à utiliser les organismes qui, à travers le pays, luttent contre les feux. Quel est l'impact des feux de forêt sur l'environnement et l'économie et comment, en sylviculture, tirer parti du feu pour un aménagement intensif des forêts, demeurent aussi la préoccupation des chercheurs.

En télédétection, la recherche veut exploiter les récentes techniques aéroportées et spatiales. À cet égard, ARIES, le système de traitement des images numériques s'avère des plus intéressants.

AMÉNAGEMENT INTENSIF DES FORÊTS. Alors que le Canada accentue ses efforts d'aménagement forestier pour répondre à une hausse prévue de la demande de bois, le rôle de ce programme d'aménagement à l'Institut prendra de plus en plus d'importance. Nul doute qu'un vaste programme de régénération intensive va nécessiter un approvisionnement sûr en semences de haute qualité des espèces voulues. En vue de constituer un meilleur stock de matériel sur pied, des recherches pour l'amélioration génétique des arbres forestiers sont grandement souhaitables. Si l'utilisation accrue de toute la partie aérienne de l'arbre (la biomasse) reste un objectif à atteindre, il faudra cependant étudier avec soin son incidence sur l'environnement; et l'on devra faire de même pour toute autre forme d'aménagement intensif. Des méthodes bio-technologiques permettant d'accroître le rendement tout en maintenant la productivité sont aussi à l'étude.

Pour mener à bien ses programmes de recherche, l'Institut dispose d'une superficie forestière de 98 km² dans la partie ouest de la réserve militaire de Petawawa. Il conserve, depuis les années 20, les registres des expériences et depuis cette même époque ne cesse d'aménager des placettes d'échantillonnage. La forêt sert également de laboratoire pour les écoles de la région. De plus, un centre d'accueil reçoit des visiteurs tout au long de l'été.

La Direction de la statistique forestière et des systèmes (DSFS), établie à l'Institut, est chargée de recueillir et de publier les données sur les ressources forestières du Canada. L'entremise du comité de l'inventaire des forêts, formé de responsables du fédéral et des provinces, facilite grandement ses rapports avec les organismes forestiers provinciaux pour améliorer et uniformiser les données sur ces ressources.

Par l'intermédiaire du programme FORSTATS, qui implique tous les centres régionaux du Service canadien des forêts, la DSFS coordonne, au sein de ce Service, la cueillette et la publication de statistiques nationales sur les forêts du Canada.

À tous les cinq ans, elle publie un inventaire des forêts du Canada. Ce rapport officiel renseigne sur les lieu, superficie, espèce et état de nos forêts. Ses nombreux contacts avec les provinces visent aussi à réunir le plus de données possibles sur les perturbations survenues en forêt: feux, récoltes, insectes, maladies, et opérations d'aménagement. Ces données sont essentielles pour planifier des stratégies à même de favoriser l'aménagement de cette ressource naturelle renouvelable et combien importante.

SYSTÈME NATIONAL D'ÉQUATIONS POUR ÉVALUER LA MASSE ANHYDRE DU PEUPLIER FAUX-TREMBLE *POPULUS TREMULOIDES* MICHX.

Rapport d'information PI-X-24F

F. Evert

Institut forestier national de Petawawa
Service canadien des forêts
Environnement Canada
1984

©Ministre des Approvisionnements et des Services, 1984
N° de catalogue Fo46-11/24-1984F
ISSN 0228-0736
ISBN 0-662-92583-1

On peut obtenir d'autres exemplaires de la présente publication du:

Centre de diffusion des publications et de l'information technique
Institut forestier national de Petawawa
Environnement Canada
Chalk River (Ontario)
K0J 1J0

N° de tél.: (613) 589-2880

This publication is also available in English under the title **A national system of equations for estimating ovendry mass of trembling aspen Populus tremuloides Michx.**

Table des matières

i	Résumé/Abstract
1	Introduction
2	Méthodes
2	Données
3	Analyse
4	Résultats
4	Comparaison des équations aux moindres carrés avec les équations aux moindres carrés pondérées
5	Applicabilité des équations régionales à toute l'aire d'extension naturelle de l'espèce
7	Préparation et essai d'un système national d'équations
8	Discussion et conclusions
11	Remerciements
12	Bibliographie
13	Annexe I: Tableaux
22	Annexe II: Notes sur l'utilisation de $s^2_{y.x}$ et R^2 pour la comparaison des équations

Figures

6	1. L'ajustement de deux équations aux moindres carrés pour l'estimation de la masse anhydre d'arbres de toute dimension.
6	2. L'ajustement de trois équations aux moindres carrés pondérées pour l'estimation de la masse anhydre d'arbres de toute dimension.

Résumé

Le rapport présente un système national unique d'équations pour évaluer, chez un peuplier faux-tremble, (Populus tremuloides Michx.) la masse anhydre de l'ensemble de sa partie aérienne et de ses éléments. Ce système est basé sur les données provenant de six régions géographiques du Canada. Comparées aux données de chaque région, les estimations de l'aggrégat de la masse anhydre de tout l'échantillon des arbres de n'importe laquelle des régions ne diffèrent pas de plus de 6 % des valeurs observées, la moitié s'en écartant de 2 %.

Abstract

This report presents a single national system of equations for estimating the aboveground ovendry mass of single trees and of their individual components for trembling aspen (Populus tremuloides Michx.). This system is based on data from six geographic regions across Canada. When applied to the sample data from individual geographic regions, estimates of aggregate ovendry mass of all sample trees in any of the six regions differed from observed values by not more than 6%, with approximately half the estimates being within 2% of observed values.

SYSTÈME NATIONAL D'ÉQUATIONS POUR ÉVALUER LA MASSE ANHYDRE DU PEUPLIER FAUX-TREMBLE POPULUS TREMULOIDES MICHX.

INTRODUCTION

Les équations standard ou à deux variables indépendantes (diamètre à hauteur de poitrine et hauteur) revêtent un aspect fondamental pour l'estimation régionale et nationale de la masse anhydre des arbres entiers et de leurs parties. Elles sont tout aussi fondamentales pour les prévisions de la croissance et de la récolte, exprimées en masse anhydre, du fait que toutes les techniques de prévision font appel en premier lieu à l'estimation de valeurs initiales ou au recensement d'un inventaire.

Il existe actuellement dix systèmes régionaux d'équations standard au Canada, permettant l'estimation de la masse anhydre du peuplier faux-tremble (Populus tremuloides Michx.), déduits de mesures directes prises sur au moins 700 arbres témoins. Chaque système régional est constitué d'équations permettant d'estimer la masse anhydre de la partie aérienne d'arbres entiers, de même que celle de chacune de ses parties: bois de la tige, écorce de la tige, branches vivantes, rameaux et feuilles. Certains systèmes régionaux ont été constitués à partir de moins de 100 arbres témoins, alors que d'autres découlent de l'étude de plus de 200 arbres témoins. Certains échantillons ont servi à l'élaboration de plus d'un système d'équations.

Il y a des raisons de penser qu'un système national d'équations intégré serait plus pratique que les nombreux systèmes régionaux actuellement en usage, et présenterait des avantages théoriques certains.

D'abord, l'échantillon combiné qui servirait à la constitution du système national intégré serait probablement plus représentatif de la grande variation des dimensions des arbres que tout autre échantillon régional. Compte tenu de ce fait, un système national unique fournirait des estimations plus précises, sur une plus grande gamme de dimensions d'arbres, que tout autre système régional, et cela dans toutes les régions. L'emploi d'un tel échantillon faciliterait en outre l'analyse de régression, parce que la variance des coefficients de régression estimés diminue avec l'élargissement du domaine des variables indépendantes (Schumacher et Chapman 1954). Dans tous les cas, un échantillon adéquat qu'il soit national ou régional pour une espèce donnée devrait inclure toute l'étendue des classes de diamètre de l'espèce, et pour chaque classe de diamètre, toute l'étendue des classes de hauteur. Les échantillons régionaux ayant ces caractéristiques sont plutôt rares compte tenu des dépenses importantes qu'il faut faire pour les constituer.

F. Evert est un chercheur de l'Institut forestier national de Petawawa.

La publication de ce manuscrit a été approuvée le 12 août 1983.

En second lieu, un système national unique faciliterait la comparaison des estimations des inventaires régionaux et des résultats des études à l'échelle nationale, par exemple ceux du programme interprovincial de fertilisation des forêts (Weetman et al. 1976), dont différentes parties sont réalisées dans un certain nombre de régions. A son crédit, il faut également ajouter qu'il faciliterait la comparaison des données sur la croissance et la récolte recueillies dans diverses régions, l'échange de ces données, et leur intégration, lorsque cela sera justifié, dans un système national unique des données sur la croissance et les récoltes.

L'objectif de la présente étude est d'intégrer tous les ensembles de données dont nous disposons pour le peuplier faux-tremble, données recueillies dans diverses régions géographiques du Canada, et d'en faire un système national unique d'équations à deux variables indépendantes. On demande à ce système de satisfaire aux exigences suivantes:

1) Mettre à notre disposition des moyens permettant d'estimer avec précision la masse anhydre d'arbres de toute dimension.
2) Faire en sorte que la somme des estimations des parties de l'arbre entier soit égale à l'estimation indépendante de l'arbre entier.
3) Donner, lorsqu'on l'applique aux échantillons de régions géographiques particulières, des estimations de la masse anhydre cumulée (Σom_i) de:

   a) tous les arbres témoins d'une région donnée, et
   b) tous les arbres témoins de chacune des trois classes dimensionnelles d'une région donnée: petite, moyenne et grande,

   qui soient comprises à l'intérieur d'un pourcentage des valeurs observées fixé à l'avance. Comme il n'existe pas de méthodes objectives pour déterminer à l'avance la valeur des pourcentages, on a fixé arbitrairement ces limites à 10 % pour tous les arbres témoins et à 15 % pour tous les arbres témoins de chacune des trois classes dimensionnelles.

4) Lorsqu'on cherche à prévoir la croissance d'un peuplement par l'intermédiaire de la prévision des variables des parties individuelles de l'équation de la masse anhydre d'un peuplement, ce système doit:

   a) rendre pratiquement faisable la prévision des variables de ces parties, et
   b) faire en sorte que la somme des parties individuelles estimées du peuplement (la masse anhydre du poids de la tige, de l'écorce de la tige, des branches vivantes, des rameaux et des feuilles) soit égale à l'estimation indépendante de la masse anhydre du peuplement entier.

MÉTHODES

Données

Les données utilisées proviennent, en premier lieu, des états imprimés d'ordinateur concernant les informations sur les arbres individuels recueillies par les six régions géographiques; en second lieu, des équations

publiées par les six études régionales. Les données de régression publiées sont illustrées au Tableau 1, alors que la répartition des renseignements sur les arbres individuels est illustrée au Tableau 2 par classe de diamètre à hauteur de poitrine de 2 cm et par classe de hauteur de 3 m, et cela pour chacune des six régions. L'écart entre le nombre d'arbres témoins indiqué au Tableau 2 et celui indiqué au Tableau 1 s'explique par un filtrage préliminaire des données de l'Ontario.

Analyse

Les solutions visant à créer un système en mesure: i) d'estimer avec précision la masse anhydre des arbres de toute dimension, et ii) de faire en sorte que la somme des estimations des parties de l'arbre soit égale à l'estimation indépendante de l'arbre entier, semblent être en corrélation. Par exemple, les équations aux moindres carrés utilisant le même modèle et ajustées à partir des mêmes données seront "additives" (Kozak 1970), alors que l'expérience nous apprend que les équations de calcul du volume des arbres donnent des estimations assez faibles pour les volumes de petits arbres (Evert 1969). Les équations aux moindres carrés pondérées ont tendance à bien estimer le volume des arbres de toutes les dimensions (Evert 1969), quoique leurs estimations ne soient pas nécessairement additives.

Pour obtenir à la fois une bonne "additivité" et de bonnes estimations de la masse anhydre des petits arbres, on a décidé d'ajuster les équations aux moindres carrés pondérées aux parties individuelles des arbres de façon cumulative plutôt que de façon séparée.

Les variables cumulées tiennent compte de la masse anhydre, exprimée en kilogramme, des parties individuelles des arbres:

bois de la tige = om(1)
bois de la tige + écorce = om(2)
bois de la tige + écorce + branches vivantes = om(3), et
bois de la tige + écorce + branches vivantes + rameaux et feuilles = om(4)

La masse des parties autres que le bois de la tige pourrait être calculée en prenant la différence des estimations appropriées obtenues de deux de ces équations. Par exemple, on pourrait connaître la masse de l'écorce de la tige en soustrayant la masse estimée obtenue par om(2) de la masse estimée obtenue par om(1).

La méthode retenue pour la pondération des écarts résiduels de la masse anhydre des arbres consiste à diviser la masse anhydre appropriée de l'arbre, om(i), par le produit de la surface terrière à hauteur de poitrine, g, par la hauteur total de l'arbre, h. Ainsi a-t-on établi la régression de ces quatres valeurs tant pour le diamètre à hauteur de poitrine que la hauteur - om(1)/gh, om(2)/gh, om(3)/gh et om(4)/gh. Les équations de la masse anhydre om(i) peuvent être obtenues dans une étape à part en multipliant les deux membres des équations ajustées à om(i)/gh par le produit gh.

Dans la première étape de l'analyse, on a déterminé si oui ou non les équations aux moindres carrés constituaient de pauvres estimateurs de la masse anhydre des petits arbres, et si oui ou non les équations aux moindres

carrés pondérées fournissaient des estimations précises pour les arbres de toutes les dimensions. Tous les modèles ont été mis à l'essai, les modèles pondérés comme les modèles non pondérés, sur la base des 327 observations disponibles pour l'Ontario, vu que cet échantillon était particulièrement bien équilibré en ce qui a trait aux dimensions des arbres (voir Tableau 2). Dans tous les modèles, on a ajusté l'équation du poids de la tige plus l'écorce, parce qu'on ne disposait pas à l'Institut forestier national de Petawawa de données sur le seul bois de la tige.

La deuxième étape de l'analyse a consisté à vérifier si oui ou non les équations établies à partir des données d'une seule région pouvaient s'appliquer à d'autres régions. Les équations de la masse anhydre standard ou à deux variables doivent donner des estimations précises pour toute l'aire d'extension naturelle parce que, même si on s'y réfère rarement d'une façon directe, ces équations fournissent les principales valeurs utilisées pour obtenir les équations locales de la masse anhydre, équations qui, elles, sont utilisées directement. La préparation des équations locales à partir des équations standard exige simplement i) la construction d'une courbe de la hauteur en fonction du diamètre à hauteur de poitrine pour le site ou le peuplement à estimer et ii) la substitution de la hauteur dans l'équation standard par sa valeur estimée en fonction du diamètre. Ainsi, dès que les équations standard sont préparées, on les utilise généralement comme base pour toutes les équations locales. Cette vérification a également été effectuée pour les équations d'estimation de la masse anhydre du bois de la tige plus l'écorce, comme dans la première étape de l'analyse.

La troisième et dernière étape a porté sur (i) la fusion de toutes les données témoins provenant des six sources régionales et sur le calcul de leur régression pour obtenir un seul système national d'équations, et (ii) la vérification de l'ajustement des équations nationales pour voir si elles estiment, autant que faire se peut, la masse anhydre du peuplier faux-tremble dans toute son aire d'extension naturelle.

Tous les calculs ont été précédés par un filtrage supplémentaire des données de base. Ainsi, les calculs définitifs ont porté sur 675 arbres témoins plutôt que sur les 695 présentés au Tableau 2.

Parce que les données concernant le bois de la tige dont dispose l'Institut sont associées à l'écorce de la tige, on a estimé le bois de la tige à partir des échantillons bois de la tige plus écorce provenant des données de l'Institut, et on a utilisé la relation bois de la tige/(bois de la tige + écorce) élaborée à partir des données restantes. Ainsi les quatre équations du système, y compris celles portant seulement sur le bois de la tige, reposent sur les mêmes 675 arbres témoins.

RÉSULTATS

Comparaison des équations aux moindres carrés avec les équations aux moindres carrés pondérées

Au Tableau 3 on peut voir deux équations aux moindres carrés et quatre équations aux moindres carrés pondérées. En outre, on peut vérifier sur les Figures 1 et 2 si oui ou non ces équations ajustent tout le domaine des

données de base qui ont servi à les déduire. La validité de l'ajustement des équations aux moindres carrés comparée à celle des équations aux moindres carrés pondérées ne peut être comparée en utilisant R^2 ou EET, en raison de l'échelle différente des écarts résiduels en cause (voir également Annexe II).

L'ajustement de ces équations a été vérifié en exprimant leurs estimations selon le pourcentage des valeurs observées, et en comparant les pourcentages obtenus à la fois par les équations aux moindres carrés et celles aux moindres carrés pondérées. L'équation (6) du Tableau 3 ne fait pas partie de la comparaison du fait qu'elle n'apporte aucune amélioration par rapport à l'équation (5), même si elle comporte une variable de plus. La somme de la masse anhydre des données de base estimées par les équations aux moindres carrés pondérées n'est pas nécessairement égale à la somme des valeurs observées comme dans le cas des estimations obtenues avec les équations non pondérées. La procédure utilisée pour corriger ce biais potentiel a été: i) de déterminer le rapport des deux sommes, c'est-à-dire la somme estimée sur la somme observée, et ii) d'utiliser la réciproque de ce rapport comme facteur de correction pour l'équation appropriée, ainsi qu'on le montre à la Figure 2.

La position des points reportés aux Figures 1 et 2 par rapport à la ligne de 100 % indique que les équations aux moindres carrés, aussi bien que celles aux moindres carrés pondérées, donnent des estimations équivalentes pour les arbres ayant un diamètre à hauteur de poitrine de 5 cm et plus, mais que pour les arbres plus petits, les équations aux moindres carrés sont des estimateurs de la masse anhydre plutôt aléatoires bien que l'échantillon de base comprenne un nombre plus élevé d'arbres de cette catégorie (voir Tableau 2). Les équations aux moindres carrés pondérées donnent en contrepartie une estimation précise des arbres de toute dimension.

Applicabilité des équations régionales à toute l'aire d'extension naturelle de l'espèce

Le Tableau 4 est une table de valeurs qui illustre la précision de cinq équations régionales à deux variables explicatives pour ce qui est de l'estimation de la masse anhydre du bois de la tige et de l'écorce du peuplier, et cela pour six échantillons régionaux prélevés à travers le Canada. Chaque ligne donne l'ajustement des cinq équations pour un échantillon régional donné, et chaque colonne, l'ajustement d'une équation régionale particulière pour les six échantillons régionaux. Toutes les estimations sont exprimées en pourcentage des valeurs observées. Plusieurs observations pertinentes à l'échantillonnage et à la préparation des équations de la masse anhydre sont évidentes à l'examen de ce tableau. Ce sont:

1) Sauf les estimations de l'équation (5) pour les échantillons de l'Alberta, de l'Ontario (1) et de l'Ontario (2), et des estimations des équations (1) et (4) pour l'échantillon de l'Institut, les cinq équations régionales donnent des estimations de la masse anhydre totale de tous les arbres pour les six échantillons régionaux dont la précision, par rapport aux valeurs observées, est de dix pour cent ou moins. Les cinq équations régionales sont basées sur des échantillons dont le nombre varie de 46 à 252 arbres témoins.

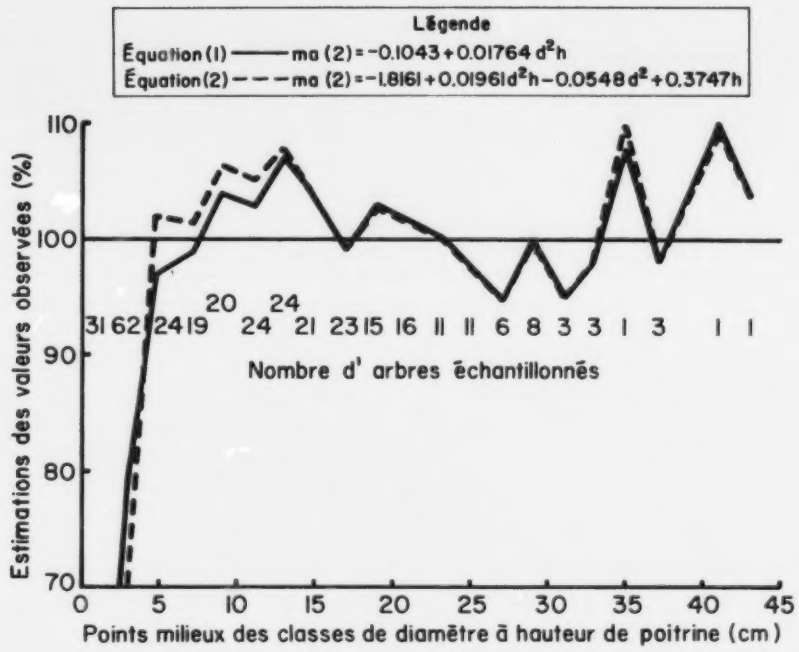


Figure 1. L'ajustement de deux équations aux moindres carrés pour l'estimation de la masse anhydre d'arbres de toute dimension.

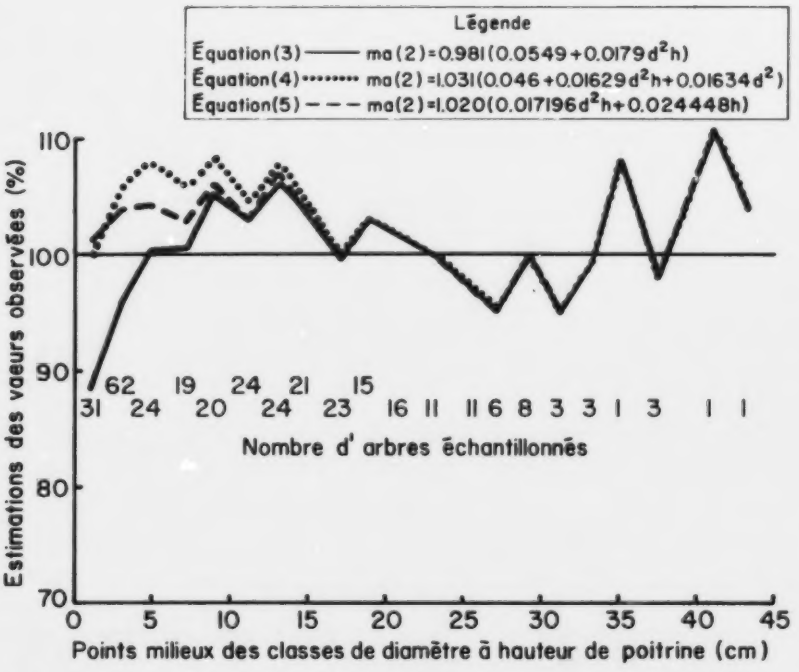


Figure 2. L'ajustement de trois équations aux moindres carrés pondérées pour l'estimation de la masse anhydre d'arbres de toute dimension.

Si l'équation (5) ne peut fournir des estimations précises des échantillons régionaux de l'Alberta, de l'Ontario (1) et de l'Ontario (2) les raisons en sont sans doute l'insuffisance de l'échantillon, seulement 46 arbres, l'éventail limité des hauteurs échantillonnées pour chaque classe de diamètre (voir Tableau 2) et l'utilisation d'un modèle de régression logarithmique:

$$\ln om = -3,0560 + 2,3536 \ln d + 0,2326 \ln h$$

ou

$$om = 0,04708 d^{2,3536} h^{0,2326}$$

Ce modèle indique, pour un diamètre donné, que la masse anhydre augmente directement avec la hauteur élevée à la puissance 0,2326. La plupart des graphiques mettent en lumière que pour un diamètre donné, la masse anhydre augmente linéairement avec la hauteur (puissance 1).

2) le défaut de l'équation régionale de l'Alberta, équation (1), qui l'empêche de fournir des estimations précises pour l'échantillon de l'Institut qui ne renferme que de petits arbres, vient renforcer la thèse exprimée précédemment montrant que les équations aux moindres carrés sont de pauvres estimateurs des volumes des petits arbres (Evert 1969) et de leur masse anhydre (voir Figure 1).

3) les estimations données par les équations (2) et (3) pour l'échantillon de l'Institut sont moins de dix pour cent par rapport aux valeurs observées en ce qui a trait à la masse anhydre totale de tous les arbres. Cependant ces équations surestiment la masse anhydre des arbres de la classe 0,1 à 2,0 cm, les estimations totales correspondent à 68,3 et 68,5 % des valeurs observées. Le modèle utilisé pour les équations (2) et (3) est:

$$om(i) = bd^2h$$

ou

$$om(i)/d^2h = b$$

Autrement dit, la grandeur om(i)/gh est présumée indépendante de la dimension de l'arbre. Toutefois le Tableau 3 met en lumière le fait que le coefficient de détermination, r^2 ou R^2, entre la grandeur om(i)/gh et la dimension de l'arbre peut atteindre 0,85 à 0,90.

Préparation et essai d'un système national d'équations

Le système national d'équations est exposé au Tableau 5, et le Tableau 6 donne les résultats d'un essai qu'on lui a fait subir.

Les quatre équations du Tableau 5 ont toutes comme premier terme $1/d^2$. Le terme 1/h est également une variable importante dans ces quatre équations, et son importance va s'accentuant de l'équation om(1) à l'équation om(4). La validité de l'ajustement des équations nationales de

om(1) et om(4) ne peut être comparée aux équations régionales correspondantes du fait que leur R^2 et leur EET respectifs utilisent des échelles décartes résiduels différentes (voir également Annexe II).

Le Tableau 6 est une table de valeurs montrant la précision des estimations de la masse anhydre obtenue, de façon cumulative, avec le système national pour le bois de la tige, l'écorce de la tige, les branches vivantes, les rameaux et les feuilles.

Chaque colonne nous fait voir l'ajustement d'une certaine équation lorsqu'elle est appliquée aux six échantillons régionaux. Toutes ces estimations sont également exprimées comme un pourcentage des valeurs observées. Aucune des estimations de la mase anhydre totale de tous les arbres témoins d'une région quelconque ne s'écarte de plus de 6 % des valeurs observées; et toutes les estimations totales de chacune des trois classes dimensionnelles des six régions ne diffèrent pas plus de 14 % des valeurs observées.

DISCUSSION ET CONCLUSIONS

À présent, il ne nous reste qu'à conclure à la lumière des preuves de ce que l'on vient de présenter, si oui ou non un système unique d'équations standard ou à deux variables satisfait les exigences précisées au début de cette étude.

D'abord, nous donne-t-il des moyens d'obtenir des estimations précises de la masse anhydre d'arbres de toute dimension? Il nous semble que la réponse à cette question est fonction de l'étendue de l'échantillonnage et de la méthode d'analyse de régression utilisée. La Figure 2 nous incite très fortement à croire que les équations s'appuyant sur un échantillon bien distribué et ajusté par la méthode des moindres carrés pondérés dont le poids est l'inverse du produit de la surface terrière par la hauteur totale de l'arbre, permettent de réaliser cette exigence.

Deuxièmement, lorsqu'on l'applique aux échantillons tirés de régions géographiques individuelles, les estimations de la masse anhydre totale de (a) tous les arbres témoins d'une région donnée et (b) de tous les arbres témoins de chacune des trois classes dimensionnelles (petite, moyenne et grande) d'une région donnée satisfont-elles les exigences de précision prescrites?

Les preuves exposées au Tableau 6, preuves résumées ci-après, nous apprennent que cette exigence est satisfaite par le système national d'équations, de la façon suivante:

Mass anhydre estimée en pourcentage des valeurs observées	Distribution des cellules (tirée du Tableau 6)			
Intervalles des classes	Petits arbres	Arbres moyens	Gros arbres	Tous les arbres
0.0 - 2.0	12	11	5	11
2.1 - 4.0	4	4	10	6
4.1 - 6.0	4	5	8	7
6.1 - 8.0	2	4	1	-
12.1 - 14.0	2	-	-	-
Tous les intervalles	24	24	24	24

L'estimation de la masse anhydre totale de: a) tous les arbres témoins de l'une quelconque des six régions diffèrent des valeurs observées par moins de 6 %, la moitié environ des estimations ne s'écartant pas plus de 2 % par rapport aux valeurs observées; et de (b) tous les arbres témoins de chacune des trois classes dimensionnelles de l'une quelconque des six régions diffèrent des valeurs observées par moins de 14 %, 97 % des estimations se trouvant à moins de 8 % des valeurs observées.

On peut rétorquer toutefois que même si les équations nationales satisfont les exigences de précision prescrites, on relève dans le Tableau 6 des preuves qui laissent entendre qu'il y aurait une sous-estimation ou une surestimation systématique dans le cas d'une certaine région. Mais quelle équation ne surestimera ou ne sous-estimera pas un échantillon particulier tiré des données qui ont servi à la déduire? Ce qui est vraiment important c'est de mesurer l'erreur et de juger si elle est acceptable.

Troisièmement, le système national unique, lorsqu'on l'utilise pour prévoir la croissance d'un peuplement:

a) est-il formé d'équations de la masse anhydre du peuplement qui permettent facilement de prévoir les variables des parties; et

b) fait-il en sorte que la somme des composantes individuelles estimées du peuplement (masse anhydre du bois de la tige, de l'écorce, des branches vivantes, des rameaux et des feuilles) soit égale à la masse anhydre estimée du peuplement entier?

Pour bien prévoir la croissance d'un peuplement, il faut disposer d'équations de la masse anhydre du peuplement qui prévoient aisément les variables du peuplement. Un système d'équations de masse anhydre de peuplement dérive du modèle de base qui a servi à estimer la masse anhydre des arbres (voir Tableau 5):

$$om = b_1 d^2 h + b_2 h + b_3 d^2$$

De la conversion algébrique de cette équation afin d'obtenir une équation de la masse anhydre du peuplement, il vient:

$$OM = \Sigma om_i$$

$$= b_1 \Sigma d_i^2 h_i + b_2 \Sigma h_i + b_3 \Sigma d_i^2$$

en décomposant en facteurs le signe $\Sigma d_i^2 h_i$, on obtient

$$\Sigma d_i^2 h_i = (\Sigma d_i^2)\ (\Sigma d_i^2 h_i)/(\Sigma d_i^2)$$

et:

$$OM = b_1\ (\Sigma d_i^2)\ (\Sigma d_i^2 h_i)/(\Sigma d_i^2) + b_2 \Sigma h_i + b_3 \Sigma d_i^2$$

où OM est la masse anhydre du peuplement pour le bois de la tige, ou le bois plus l'écorce, etc. dont on veut connaître une estimation;
Σd_i^2 est la somme des diamètres au carré;
$(\Sigma d_i^2 h_i)/(\Sigma d_i^2)$ est la hauteur moyenne pondérée avec la surface terrière, ou la hauteur de Lorey (h_i), basée sur les hauteurs estimatives de tous les arbres dont on veut connaître une estimation de la masse anhydre; et
Σh_i est la somme des hauteurs estimées de tous les arbres dont on veut connaître une estimation de la masse anhydre.

La somme des hauteurs estimées de tous les arbres peut également être exprimée sous forme du produit du nombre d'arbres, N, par la moyenne arithmétique de la hauteur de tous les arbres, $\bar{h}$,:

$$\Sigma h_i = N(\Sigma h_i/N) = N\bar{h}$$

c'est une façon plus simple de prévoir la croissance d'un peuplement. En outre, la somme des diamètres au carré de tous les arbres peut également être exprimée sous la forme du produit du nombre d'arbres, N, par le carré du diamètre moyen, d_g^2:

$$\Sigma d_i^2 = N d_g^2$$

Ainsi:

$$OM = N d_g^2 (b_1 h_L + b_3) + b_2 N\bar{h}$$

La réponse à la troisième question semble donc être oui. Toutes les variables des éléments individuels de l'équation de la masse anhydre du peuplement, exposées ci-dessus, ont servi à la prévision du volume du peuplement et de sa croissance. Il apparaît donc clairement qu'on peut l'utiliser pour prévoir la masse anhydre du peuplement et la croissance de ce peuplement.

La méthode algébrique de déduction des équations de la masse anhydre du peuplement à partir des équations de la masse anhydre d'un arbre nous assure que la somme de la masse anhydre estimée des éléments individuels du peuplement égale la masse anhydre estimée du peuplement entier.

Suite aux preuves que l'on vient de donner, on ne croit pas errer en tirant les conclusions suivantes:

1) Pour préparer un système d'équations de la masse anhydre valable tant pour des applications régionales que nationales, on devrait:

   a) s'appuyer sur un échantillon qui inclut toutes les classes de diamètre d'une espèce et, dans chaque classe de diamètre, tout l'intervalle des classes de hauteur. L'éventail dimensionnel des arbres témoins semble être plus important que leur répartition géographique;

   b) utiliser comme méthode d'ajustement la méthode des moindres carrés pondérés plutôt que la seule méthode aux moindres carrés, en utilisant comme poids la réciproque du produit de la surface terrière de l'arbre par la hauteur totale.

2) Pour être certain que les équations sont additives lorsqu'on les ajuste par la méthode aux moindres carrés pondérés, l'ajustement doit être effectué d'une façon cumulative pour les parties individuelles des arbres plutôt que séparément. Les conditions prescrites par Kozak (1970) et données ci-après doivent également être satisfaites:

   a) on doit ajuster exactement le même modèle à toutes les équations;

   b) toute transformation de la variable dépendante doit être linéaire;

   c) toutes les équations d'un système donné doivent être ajustées à partir des mêmes observations.

3) Afin de s'assurer que les équations de la masse anhydre des arbres manifestent un bon ajustement et sont additives, on doit déduire algébriquement les équations de la masse anhydre du peuplement à partir des équations de la masse anhydre des arbres.

À titre de recommandation générale, il nous semble important d'examiner la possibilité d'intégrer en un seul système national tous les autres systèmes régionaux actuellement en usage et qui portent sur des espèces ayant une grande aire d'extension naturelle.

REMERCIEMENTS

Les informations sur les arbres qui ont servi à cette étude proviennent des états imprimés d'ordinateur des centres de recherche du Service canadien des forêts. L'auteur tient à exprimer tous ses remerciements aux personnes qui lui ont fait parvenir ces données.

CRF Maritimes	M.F. Ker
CRF Terre-Neuve	R.S. van Nostrand
CRF du Nord	I.E. Bella et T. Singh
Institut forestier national de Petawawa	I.S. Alemdag

Les calculs sont l'oeuvre de M. W.A. Kean, et le traçage des figures effectué par M. D. McGuire. L'auteur les remercie de l'aide qu'ils lui ont apportée.

BIBLIOGRAPHIE

Alemdag, I.S.; Horton, K.W. 1981. Single-tree equations for estimating biomass of trembling aspen, largetooth aspen and white birch in Ontario. For. Chron. 57(4): 169-173.

Berry, A.B.; Stiell, W.M. 1978. Effect of rotation length on productivity of aspen sucker stands. For. Chron. 54(5): 265-267.

Evert, F. 1969. Use of form factor in tree-volume estimation. J. For. 67(2): 126-128.

Johnstone, W.D.; Peterson, E.B. 1980. Above-ground component weights in Alberta Populus stands. Can. Dep. Environ., Can. For. Serv., North. For. Res. Cent., Inf. Rep. NOR-X-226. vi-18 p.

Ker, M.F. 1980. Tree biomass equations for ten major species in Cumberland county, Nova Scotia. Can. Dep. Environ., Can. For. Serv., Maritimes For. Res. Cent., Inf. Rep. M-X-108. iii+26 p.

Kozak, A. 1970. Methods of ensuring additivity of biomass components by regression analysis. For. Chron. 46(5): 402-404.

Lavigne, M.B.; van Nostrand, R.J. 1981. Biomass equations for six tree species in central Newfoundland. Can. Dep. Environ., Can. For. Serv., Nfld. For. Res. Cent., Inf. Rep. N-X-199. v+51 p.

Schumacher, F.X.; Chapman, R.A. 1954. Sampling methods in forestry and range management. P. 137, Bulletin 7, revised, School of Forestry, Duke University, Durham, North Carolina. 221 p.

Weetman, G.F.; Krause, H.H.; Koller, E. 1976. Interprovincial forest fertilization program. Results of five-year growth remeasurements in thirty installations: fertilized in 1969, remeasured in 1974. Can. Dep. Environ., Can. For. Serv., For. Techn. Rep. 16, v+34 p.

ANNEXE I :
TABLEAUX

Tableau 1. Données de régression sur la prévision de la masse anhydre, om(i) du peuplier faux-tremble dans six régions géographiques

a) Ontario

(i) Régions forestières Grands Lacs - Saint-Laurent (Alemdag & Horton 1981)
Modèle om(kg) = bd^2h (cm) (m)
n = 128, éventail des diamètres à hauteur de poitrine de 5,2 à 43,5 cm

	b
om(1)[1]	0,014483
om(2)[2]	0,017939
om(3)[3]	0,020610
om(4)[4]	0,021087

(ii) Région forestière boréale (Alemdag & Horton 1981)
Modèle om(kg) = bd^2h (cm) (m)
n = 96, éventail des diamètres à hauteur de poitrine de 5,2 à 35,4 cm

	b
om(1)	0,014748
om(2)	0,017496
om(3)	0,019690
om(4)	0,020259

(iii) Institut forestier national de Petawawa (Berry & Stiell 1978)
Modèle om(kg) = $b_0 + b_1d^2h + b_2(d^2h)^2$ (cm) (m)
n et éventail des diamètres à hauteur de poitrine non communiqués

	b_0	b_1	b_2
om(3)	0,0992	0,02175	$-7,01 \times 10^{-6}$

b) Alberta (Johnstone & Peterson 1980)
Modèle om(kg) = $b_0 + b_1d + b_2d^2 + b_3d^3 + b_4h + b_5d^2h$ (cm) (m)
n = 254, éventail des diamètres à hauteur de poitrine de 2,0 à 31,5 cm

	b_0	b_1	b_2	b_3	b_4	b_5
om(1)	1,493	0,2384	-0,0046	-0,0004	-0,3040	0,0144
om(2)	1,618	0,3110	0,0178	-0,0005	-0,3916	0,0167
om(3)	-1,367	2,4238	-0,1551	0,0043	-0,8636	0,0183
om(4)	-1,316	2,5077	-0,1566	0,0045	-0,9072	0,0184

c) Terre-Neuve (Lavigne & van Nostrand 1981)
Modèle om(kg) = $b_0 + b_1d^2h$ (cm) (m)
n = 70, éventail des diamètres à hauteur de poitrine de 2,2 à 44,7 cm

	b_0	b_1
om(2)	0,2420	0,01691
om(3)	0,1927	0,01933
om(4)	0,5497	0,01987

Tableau 1. (suite)

d) Nouvelle-Écosse (Ker 1980)
Modèle ln om (kg) = b_0 + b_1 ln d + b_2 ln h (cm) (m)
n = 46, éventail des diamètres - hauteur de poitrine de 1,8 à 33,3 cm

	b_0	b_1	b_2
om(1)	-3,3596	2,3670	0,2548
om(2)	-3,0560	2,3536	0,2326
om(4)	-2,6224	2,4827	-

[1] bois de la tige
[2] bois de la tige plus écorce
[3] bois de la tige plus écorce plus branches vivantes
[4] bois de la tige plus écorce plus branches vivantes plus rameaux et feuilles

Tableau 2. Nombre d'arbres témoins répartis par endroit, par classe de diamètre à hauteur de poitrine de 2 cm, et par classe de hauteur de 3 cm

Éventail des diamètres à hauteur de poitrine (cm)	Régions forestières Ontario-Grands Lacs, St-Laurent									Régions forestières Ontario-Borden									Ontario-IFNP				Alberta									
	Éventail des hauteurs (m)																														
	4,6-7,5	7,6-10,5	10,6-13,5	13,6-16,5	16,6-19,5	19,6-22,5	22,6-25,5	25,6-28,5	Toutes les hauteurs	7,6-10,5	10,6-13,5	13,6-16,5	16,6-19,5	19,6-22,5	22,6-25,5	25,6-28,5	Toutes les hauteurs	1,6-4,5	4,6-7,5	7,6-10,5	Toutes les hauteurs	1,6-4,5	4,6-7,5	7,6-10,5	10,6-13,5	13,6-16,5	16,6-19,5	19,6-22,5	22,6-25,5	25,6-28,5	Toutes les hauteurs
0,1 - 2,0																		28	3		31	1									1
2,1 - 4,0																		17	45		62	1									10
4,1 - 6,0		5							5	2							2		10	7	17		9								20
6,1 - 8,0	3	4	4						11	5	3						8						9	11	6						22
8,1 - 10,0		2	11	2					15	1	3	1					5						7	9	26	6	2	1			49
10,1 - 12,0			5	8	1				14	-	2	6	2				10						14	3	12	8	8	1			35
12,1 - 14,0			-	7	7	2			15	1	-	5	3				9						12	-	4	13	1	-			30
14,1 - 16,0			1	1	8	3	1		11			4	4	1			10						4	-	5	8	1	2			19
16,1 - 18,0			-	-	5	2	-		13			1	3	-			7							1	-	11	1	-			17
18,1 - 20,0			1	1	1	2	5		8				2	1			5									7	1	4	1		13
20,1 - 22,0						-	5		9				1	1			3									7	1	2	1		11
22,1 - 24,0						1	6		5					1			2									2	1	-	1		7
24,1 - 26,0							2		6								1									1	1	2	1		6
26,1 - 28,0							1		3								1									1	1	-	2		6
28,1 - 30,0							-	1	2								1										1	2	1		3
30,1 - 32,0							1	-	1								1											1	1		3
32,1 - 34,0							-	1	1								1														
34,1 - 36,0							1	-	2								1														
36,1 - 38,0							-	2	2								1														
38,1 - 40,0							1		1								1														
40,1 - 42,0							1		1								1														
42,1 - 44,0																															
44,1 - 46,0																															
Tout l'éventail	3	11	22	20	29	11	24	4	124	9	8	17	25	26	6	2	93	45	58	7	110	2	55	24	53	65	21	13	13	6	252

Tableau 2. (suite)

Terre-Neuve										Nouvelle-Écosse							Tous les endroits									
1,6-4,5	4,6-7,5	7,6-10,5	10,6-13,5	13,6-16,5	16,6-19,5	19,6-22,5	22,6-25,5	Toutes les hauteurs		1,6-4,5	4,6-7,5	7,6-10,5	10,6-13,5	13,6-16,5	16,6-19,5	Toutes les hauteurs		1,6-4,5	4,6-7,5	7,6-10,5	10,6-13,5	13,6-16,5	16,6-19,5	19,6-22,5	22,6-25,5	25,6-28,5	Toutes les hauteurs
2	3	1	1	1	2	2	1	5		1	3					1		30	3	19	13	10	5	1	2	1	33
1	3	-	1	-	3	3	-	5		3	1	3	3	1	1	3		23	57	21	41	22	18	2	5	-	80
	3	3	-	3	3	4	1	5		1	1	1	1	2	2	3		1	32	10	23	29	14	1	5	1	52
		2	2	3	3	5	-	5					1	1	1	1			13	2	4	19	14	11	7	1	47
		1	-	1	4	2	1	4				1	-	2	4	4			15	2	18	11	8	7	8	2	76
		-	2	2	1	-	-	3				-	1	1	1	2			13	9	1	10	8	10	5	1	66
		3	1	-	2	1	1	4				1	-	2	3	2			4	1	2	6	5	6	4	2	59
		2	-	3	2	-	-	4				1	1	1	2	2				1	1	2	3	8	3	1	49
		1	1	2	1	1	1	4				2	2	1	2	2				1	-	1	2	5	5	1	48
		1	1	1	2	1	1	4				1	1	2	1	2				1	1	1	2	4	1	2	40
		1	1	1	2	1	1	4				1	1	2	1	2				1	1	1	2	2	1	1	32
		1	1	1	2	1	1	4				1	1	2	1	2				1	1	1	2	1	-	1	30
		1	1	1	2	1	1	4				1	1	2	1	2				1	1	1	2	1	2	1	23
		1	1	1	2	1	1	4				1	1	2	1	2				1	1	1	2	1	1	2	20
		1	1	1	2	1	1	4				1	1	2	1	2				1	1	1	2	1	2	1	17
		1	1	1	2	1	1	4				1	1	2	1	2				1	1	1	2	1	1	2	10
		1	1	1	2	1	1	4				1	1	2	1	2				1	1	1	2	1	2	1	5
		1	1	1	2	1	1	4				1	1	2	1	2				1	1	1	2	1	1	2	1
		1	1	1	2	1	1	4				1	1	2	1	2				1	1	1	2	1	2	1	3
		1	1	1	2	1	1	4				1	1	2	1	2				1	1	1	2	1	1	2	-
		1	1	1	2	1	1	4				1	1	2	1	2				1	1	1	2	1	2	1	2
		1	1	1	2	1	1	4				1	1	2	1	2				1	1	1	2	1	1	2	1
		1	1	1	2	1	1	4				1	1	2	1	2				1	1	1	2	1	2	1	1
3	9	7	4	17	23	5	2	70	4	5	5	9	10	13	46	54	137	65	105	130	94	54	44	12	695	

Tableau 3. Équations d'estimation de la masse anhydre, constituées à partir de 327 observations faites en Ontario

N° d'équation	Équation	r^2 ou R^2	EET %
	Équations aux moindres carrés		
(1)	$om(2)^* = -0,1043 + 0,01764d^2h$	0,985	19,4
(2)	$om(2) = -1,8161 + 0,01961d^2h - 0,0548d^2 + 0,3747h$	0,985	19,4
	Équations aux moindres carrés pondérées		
(3)	$om(2)/gh = 228,417 + 699,421(1/d^2h)$	0,847	12,6
	ou		
	$om(2) = 0,0549 + 0,01794d^2h$	-	-
(4)	$om(2)/gh = 207,447 + 587,70(1/d^2h) + 208,05(1/h)$	0,876	11,4
	ou		
	$om(2) = 0,0460 + 0,01629d^2h + 0,01634d^2$	-	-
(5)	$om(2)/gh = 218,943 + 311,286(1/d^2)$	0,898	10,3
	ou		
	$om(2) = 0,017196d^2h + 0,024448h$	-	-
(6)	$om(2)/gh = 218,461 + 309,837(1/d^2) + 5,3084(1/h)$	0,898	10,3
	ou		
	$om(2) = 0,01716d^2h + 0,02433h + 0,0004169d^2$	-	-

*om(2) = masse anhydre du bois de la tige et de l'écorce

Tableau 4. L'ajustement de cinq équations régionales d'estimation de la masse anhydre à six échantillons régionaux du Canada

Emplacement de l'échantillon	Éventail des diamètres à hauteur de poitrine(cm)	Nombre d'arbres témoins	Équations* (1)	(2)	(3)	(4)	(5)**
			Masse anhydre estmée du bois de la tige et de l'écorce exprimée en pourcentage des valeurs observées				
Alberta	0,1 - 10,0	102	100,1	96,7	94,4	93,5	91,0
	10,1 - 16,0	84	100,7	102,7	100,2	97,4	87,0
	16,1 - 32,0	66	100,8	105,7	103,1	99,8	84,6
	Tous	252	100,7	104,4	101,9	98,8	85,6
Ontario (1)	4,1 - 12,0	45	99,5	101,9	98,4	97,4	84,7
Régions forestières	12,1 - 18,0	39	97,5	102,4	99,9	96,9	76,3
des Grands Lacs et	18,1 - 44,0	40	95,3	101,2	98,8	95,5	76,6
du St-Laurent	Tous	124	95,9	101,5	99,0	95,9	77,0
Ontario (2)	4,1 - 14,0	34	105,1	110,1	107,4	104,7	85,4
Région forestière	14,1 - 20,0	27	100,8	105,6	103,0	99,8	80,2
boréale	20,1 - 36,0	32	95,3	100,6	98,1	94,9	80,2
	Tous	93	97,1	102,3	99,8	96,6	80,6
Ontario (3)	0,1 - 2,0	31	368,1	68,3	66,5	159,5	78,8
IFNP	2,1 - 4,0	62	146,4	92,0	89,8	111,6	96,8
	4,1 - 6,0	17	115,6	101,0	98,5	103,3	101,1
	Tous	110	147,9	94,3	92,0	112,1	97,4
Terre-Neuve	2,1 - 14,0	25	107,3	102,6	100,1	98,5	98,0
Centre	14,1 - 22,0	21	97,3	100,5	98,1	95,0	85,4
	22,1 - 46,0	24	102,4	107,8	105,2	101,7	96,4
	Tous	70	101,4	105,9	103,3	99,9	93,8
Nouvelle-Écosse	0,1 - 12,0	16	95,9	85,2	83,1	82,6	92,3
comté de	12,1 - 22,0	14	99,3	100,9	98,3	95,3	94,5
Cumberland	22,1 - 34,0	16	106,0	110,5	107,8	104,3	105,9
	Tous	46	103,9	107,2	104,5	101,3	102,6

*(1) $om(2) = 1,6176 + 0,311d + 0,0178d^2 - 0,0005d^3 - 0,3916h + 0,0167d^2h$ (Alberta)
(2) $om(2) = 0,01794d^2h$ (Ontario (1))
(3) $om(2) = 0,01750d^2h$ (Ontario (2))
(4) $om(2) = 0,2420 + 0,01691d^2h$ (Terre-Neuve)
(5) $\ln om(2) = -3,0560 + 2,3536 \ln d + 0,2326 \ln h$ (Nouvelle-Écosse)
**On a utilisé l'équation (5) telle quelle sans appliquer le facteur de correction (1,01) suggéré par les auteurs, parce qu'à toute fin pratique il ne change pas les estimations obtenues.

Tableau 5. Données de régression sur la production de la masse anhydre du peuplier faux-tremble, établies à partir de 675 observations provenant de six échantillons régionaux de/par le Canada

	Équation	R^2	EET %
Bois de la tige	$om(1)/gh = 181,987 + 181,911(1/d^2) - 39,5058(1/h)$	0,559	11,0
	ou		
	$om(1) = 0,014293d^2h + 0,014287h - 0,0003103d^2$	-	-
Bois de la tige + l'écorce	$om(2)/gh = 213,425 + 280,854(1/d^2) + 95,0715(1/h)$	0,735	11,1
	ou		
	$om(2) = 0,01676d^2h + 0,022058h + 0,0074669d^2$	-	-
Bois de la tige + l'écorce + les branches vivantes	$om(3)/gh = 225,672 + 246,244(1/d^2) + 303,008(1/h)$	0,704	12,4
	ou		
	$om(3) = 0,017724d^2h + 0,01934h + 0,023798d^2$	-	-
Bois de la tige + l'écorce + les branches vivantes + les rameaux et les feuilles	$om(4)/gh = 225,033 + 244,864(1/d^2) + 434,416(1/h)$	0,708	13,5
	ou		
	$om(4) = 0,01767d^2h + 0,01923h + 0,034119d^2$	-	-

Tableau 6. L'ajustement de quatres équations nationales d'estimation de la masse anhydre du peuplier faux-tremble à six échantillons régionaux de par le Canada

Emplacement de l'échantillon	Éventail des diamètres à hauteur de poitrine	Nombre d'arbres témoins	Équations* (1)	(2)	(3)	(4)
			Masse anhydre exprimée en pourcentage des valeurs observées			
Alberta	2,1 - 10,0	90	99,8	97,9	103,9	106,0
	10,1 - 16,0	82	102,7	100,3	105,2	107,0
	16,1 - 32,0	67	103,2	101,5	101,9	102,9
	Tous	239	102,9	101,0	102,7	104,0
Ontario (1)	4,1 - 12,0	45	99,0	100,3	101,8	98,0
Régions forestières	12,1 - 18,0	39	99,7	98,9	102,6	101,5
des Grands Lacs	18,1 - 44,0	40	98,6	97,5	95,6	95,1
St-Laurent	Tous	124	98,9	97,9	97,2	96,4
Ontario (2)	4,1 - 14,0	34	101,7	107,4	114,0	113,4
Région forestière	14,1 - 20,0	27	98,6	101,8	106,3	106,1
boréale	20,1 - 36,0	32	94,1	96,4	96,5	95,9
	Tous	93	95,6	98,3	99,7	99,2
Ontario (3)	0,1 - 2,0	30	100,9	101,9	105,4	104,9
IFNP	2,1 - 4,0	62	104,9	105,2	105,8	106,3
	4,1 - 6,0	17	105,8	105,8	104,1	103,9
	Tous	109	105,0	105,3	105,0	105,4
Terre-Neuve	2,1 - 14,0	23	98,7	101,5	106,9	102,6
Centre	14,1 - 22,0	21	96,5	97,1	99,5	99,8
	22,1 - 46,0	24	102,7	103,5	101,6	101,3
	Tous	68	101,1	101,9	101,3	101,0
Nouvelle-Écosse	2,1 - 14,0	14	100,0	101,1	103,9	104,8
comté de Cumberland	14,1 - 22,0	12	104,5	100,5	99,2	100,5
	22,1 - 34,0	16	105,2	106,4	97,8	98,9
	Tous	42	104,8	104,9	98,3	99,4

*(1) $om(1) = (0,014293d^2h + 0,014287h - 0,0003103d^2) \times 1,002$
(2) $om(2) = (0,01676d^2h + 0,022058h + 0,0074669d^2) \times 1,003$
(3) $om(3) = (0,017724d^2h + 0,01934h + 0,023798d^2) \times 1,028$
(4) $om(4) = (0,01767d^2h + 0,01923h + 0,034119d^2) \times 1,028$

ANNEXE II:

Notes sur l'utilisation de $s^2_{y.x}$ et R^2 pour la comparaison des équations

Annexe II. Notes sur l'utilisation de $s^2_{y.x}$ et R^2 pour la comparaison des équations

D.A. MacLeod
Direction générale de l'informatique et de la statistique appliquée
Place Vincent Massey, Hull, Québec

Lorsqu'en régression linéaire, on ajuste une équation de la forme:

$$Y = a + b_1X_1 + b_2X_2 + \ldots + b_mX_m$$

à un ensemble de valeurs, les coefficients a, b_1, b_2 ..., b_m obtenus sont d'une certaine façon les meilleurs, c'est-à-dire ceux qui minimisent la somme des carrés des écarts résiduels (S.C.É.R.). D'autres grandeurs que l'on peut calculer sont le carré moyen des écarts résiduels $s^2_{y.x}$, donné par l'expression

$$s^2_{y.x} = \text{S.C.É.R.}/(N-M) \text{ où } n = \text{nombre de données}$$

et le coefficient de corrélation multiple R^2, donné par:

$$R^2 = 1 - \text{S.C.É.R.}/\text{S.C.E.}(Y)$$

où S.C.E.(Y) est la somme des carrés des écarts $Y_i - \overline{Y}$. Puisque la S.C.É.R. est toujours inférieure à la S.C.E.(Y) et représente la partie de S.C.E.(Y) qui n'est pas expliquée par la régression, R^2 peut être considéré comme la fraction de S.C.E.(Y) qui est expliqué.

Envisageons R^2 sous une autre optique: si on a utilisé Y = a pour l'ajustement des données, la valeur obtenue pour a sera $\overline{Y}$, et les écarts $Y_i - \overline{Y}$ seront les écarts résiduels. On voit que R^2 représente la fraction de la somme des carrés des écarts qui est expliquée lorsque les termes ($b_1X_1 + b_2X_2 + \ldots b_mX_m$) sont additionés à l'équation Y = a.

Posons que om, d et h représentent respectivement la biomasse, le diamètre et la hauteur d'un arbre. Chacune des trois équations analysées dans la présente étude:

$$om = a + b(d^2h) \quad (1)$$

$$om = b(d^2h) \quad (2)$$

$$om/d^2h = a + b/d^2 \quad (3)$$

découle d'hypothèses différentes. L'une des différences dans les hypothèses réside dans la forme de l'équation: l'équation (1) établit un rapport linéaire direct entre om et d^2h, l'équation (2) est une ligne droite qui passe par l'origine (c'est-à-dire om = 0 lorsque $d^2h = 0$) et l'équation (3), lorsqu'on multiplie chaque membre par d^2h, donne l'équation plus complexe $om = a(d^2h) + bh$ qui correspond à une courbe passant également par l'origine. Une autre différence est l'échelle des écarts résiduels: dans l'équation (1) et l'équation (2), il a été présumé que les écarts résiduels de om doivent être minimisés, alors que dans l'équation (3), ce sont ceux de om/d^2h qui doivent être minimisés.

Si on veut comparer la validité de l'ajustement de ces équations à un ensemble de données, en utilisant soit $s^2_{y.x}$ ou R^2, des difficultés surgiront à cause des différentes hypothèses. On peut comparer directement les équations (1) et (2) à l'aide de $s^2_{y.x}$ (sous réserve que la formule exacte S.C.É.R./(n-1) est utilisée pour l'équation (2)), alors qu'aucune de celles-ci ne peut être comparée à l'équation (3) du fait que les écarts résiduels sont exprimés dans une échelle différente. Il est possible de transformer l'équation (3) et d'obtenir les écarts résiduels sur l'échelle om, et ensuite à partir de ces écarts obtenir un $s^2_{y.x}$ transformé. Cependant cette façon de faire désavantage l'équation (3) car celle-ci a été obtenue en minimisant les écarts résiduels sur l'échelle om/d^2h, et non sur celle de om.

Quand l'ajustement à un ensemble de données doit être comparé avec l'ajustement à un ensemble différent, une autre difficulté vient s'ajouter

si, comme c'est tout probable, les écarts résiduels tendent à augmenter pour les arbres de plus grande dimension. Dans le cas présent $s^2_{y.x}$ représente la moyenne totale et non la valeur de chaque point de l'éventail de d ou h. Pour que le $s^2_{y.x}$ d'un ensemble de données soit comparable au carré moyen des écarts résiduels d'un autre ensemble, il faut que les ensembles de données aient le même éventail de d et h.

L'utilisation de $s^2_{y.x}$ pour effectuer des comparaisons soulève, comme on l'a déjà vu, des problèmes, mais l'utilisation des R^2 en soulève de bien plus graves. En fait, il n'est pas du tout possible d'effectuer de bonnes comparaisons entre les équations (1), (2) ou (3) à l'aide des valeurs de R^2 fournies par les programmes informatiques standard.

La valeur de R^2 est valide pour l'équation (1) et elle représente l'amélioration lorsqu'on ajoute $b(d^2h)$ à om = a. Dans le cas de l'équation (2), la valeur de a est égale à zéro. La plupart des programmes informatiques ne font pas cette correction et calculent simplement l'amélioration lorsqu'on ajoute $b(d^2h)$ à om = 0. Comme l'équation om = 0 n'ajuste pas du tout les données (contrairement à om = a), l'amélioration due à l'addition de $b(d^2h)$ est énorme. La valeur de R^2 résultant est donc artificiellement grande.

En revanche, R^2 pour l'équation (3) est artificiellement trop petite. Elle représente l'amélioration lorsqu'on ajoute b/d^2 à $om/d^2h = a$. En multipliant les deux membres par d^2h, on voit que l'amélioration est celle obtenue en ajoutant bh à om = (d^2h). Mais l'équation om = $a(d^2h)$ est déjà un ajustement très acceptable, de sorte que l'amélioration résultant de l'addition de bh n'est pas très importante. Ceci donne donc un R^2 faible même si l'ajustement est très bon. Un autre objet de difficulté est le fait

que l'amélioration est calculée d'après l'échelle om/d^2h, et non d'après l'échelle de om comme pour les équations (1) et (2).

Une solution possible, même si ce n'est pas la solution idéale, est de modifier R^2 afin d'en faire une grandeur plus comparable. On doit donc d'abord décider sur quelle échelle porteront les comparaisons. Pour plus de simplicité, supposons que l'on choisit l'échelle om. Ensuite on obtient la somme des carrés des écarts résiduels (S.C.É.R.) dans cette échelle pour chaque équation, en tranformant le cas échéant les équations, et on définit un R^2 modifié comme suit:

$$R^2(\text{mod}) = 1\text{-S.C.É.R./S.C.E.(om)}$$

où S.C.E. (om) est la somme des carrés des écarts $om_i - \overline{om}$. Cette modification a pour effet de comparer chaque équation (après sa transformation le cas échéant afin que om soit seul dans le membre de droite) avec l'équation om = a. Pour l'équation (1), R^2(mod) correspond simplement à la valeur courante de R^2. Pour les équations (2) et (3), il faut procéder à des calculs supplémentaires pour obtenir R^2(mod).

Finalement, il est bon de faire remarquer que R^2(mod) compare les équations à l'échelle des arbres individuels à l'aide des écarts $(om_i - \overline{om})$. En pratique, toutfois, l'utilisation définitive des équations est d'obtenir la biomasse de peuplements plutôt que celle d'arbres individuels. Il n'y a aucune certitude que l'équation qui s'ajuste le mieux aux données sur un arbre particulier sera cella qui s'ajuste le mieux au peuplement. La seule façon de déterminer cela est d'appliquer les équations des peuplements à des valeurs de biomasse connues, comme on le fait dans ce rapport.